



LOYOLA COLLEGE (AUTONOMOUS), CHENNAI – 600 034

M.Sc. DEGREE EXAMINATION – COMPUTER SCIENCE

THIRD SEMESTER – APRIL 2018

16PCS3MC01- BIG DATA ANALYTICS

Date: 24-04-2018
Time: 09:00-12:00

Dept. No.

Max. : 100 Marks

PART-A

Answer ALL the questions

10 x 2=20

- 1) What is structured and unstructured data?
- 2) Give reasons to sample data.
- 3) Define data stream.
- 4) What is DSMS?
- 5) Why is Hadoop used?
- 6) What is MapReduce?
- 7) Mention the ways to set up Hadoop clusters.
- 8) What is IBM BigInsights?
- 9) List Hadoop related projects.
- 10) What is Hue?

PART - B

Answer ALL the questions

5 x 8=40

- 11) a) Explain the challenges of a conventional system for analyzing data.
(or)
b) Describe five key approaches to analyse big data.
- 12) a) Explain bloom filter.
(or)
b) Give a brief description on the types of windows in data stream analysis.
- 13) a) Explain HDFS architecture
(or)
b) Describe MapReduce shuffle function.
- 14) a) Explain setting up Hadoop clusters from the scratch..
(or)
b) Elucidate Apache Ambari.
- 15) a) Explain HBase.
(or)
b) Explain pig.

PART- C

Answer ANY TWO questions

2 x 10 = 20

- 16) a) Describe the characteristics of big data. (10)
b) Give explanation on data analytics life cycle. (10)
- 17) Explain the following:
a) Flajolet-martin Approach. (10)
b) Role of JobTracker and TaskTracker with diagram. (10)
- 18) Elucidate the following:
a) Working with files in Hadoop using imdemocloud. (10)
b) Hive. (10)

\$\$\$\$\$\$\$\$